

A possible classification of nonequilibrium steady states

R. K. P. Zia and B. Schmittmann

Center for Stochastic Processes in Science and Engineering and
Department of Physics, Virginia Tech, Blacksburg, VA 24061-0435 USA

Abstract. We propose a general classification of nonequilibrium steady states in terms of their stationary probability distribution and the associated probability currents. The stationary probabilities can be represented graph-theoretically as directed labelled trees; closing a single loop in such a graph leads to a representation of probability currents. This classification allows us to identify all choices of transition rates, based on a master equation, which generate the same nonequilibrium steady state. We explore the implications of this freedom, e.g., for entropy production.

PACS numbers: 05.70.Ln, 2.50.Ga, 02.70.Rr, 05.50.+q

E-mail: rkpzia@vt.edu, schmittm@vt.edu

Submitted to: *J. Phys. A: Math. Gen.*

Introduction. One of the greatest successes of statistical mechanics was forging the fundamental link between microscopic interactions and macroscopic behavior for interacting many-body systems in equilibrium. Boltzmann and Gibbs established the general framework which allows us to compute – at least in principle – any macroscopic observable of interest. Labeling the microscopic states (“configurations”, \mathcal{C}) of the system and establishing a form for the internal energy $\mathcal{H}(\mathcal{C})$ associated with each \mathcal{C} , macroscopic observables can be expressed as statistical averages, $\langle A \rangle = \sum_{\mathcal{C}} A(\mathcal{C}) P^{eq}(\mathcal{C})$, with the appropriate equilibrium distribution $P^{eq}(\mathcal{C})$.

Given that a real system continuously undergoes transitions from one configuration to another, it is quite remarkable that $P^{eq}(\mathcal{C})$ can be determined without explicit recourse to a time-dependent distribution, $P(\mathcal{C}; t)$. At the root of this enormous simplification lies the property of *detailed balance*. Related to microscopic reversibility, a system evolving according to a dynamics with this property will eventually settle in a stationary state in which the *net* probability current between *any pair* of configurations vanishes. As a result, $P^{eq}(\mathcal{C}) \equiv \lim_{t \rightarrow \infty} P(\mathcal{C}; t)$ can be expressed in terms of *ratios* of the (dynamic) transition rates between configurations, and the long-time limit remains invariant under any modification of the dynamics which preserves these *ratios*. Indeed, Monte Carlo simulation studies of equilibrium systems rely heavily on this property.

In summary, systems in thermal equilibrium are fundamentally well understood, including their dynamical representations. In stark contrast, a comparable theoretical framework is still sorely lacking for systems far from thermal equilibrium. Even the

simplest generalizations of thermal equilibrium, namely, non-equilibrium steady states (NESS), are currently analyzed case-by-case. Thus, much effort is directed at simple models, maintained far from equilibrium by imposing external driving forces, with the goal of identifying some generic classes of NESS [1, 2]. Typically, these models are specified by a set of transition rates (motivated by physical considerations), so that the master equation provides a natural framework for analysis. A key feature of far-from-equilibrium dynamics, these rates *violate* detailed balance, resulting in *non-vanishing* probability currents for the final, time-independent NESS. In general, its stationary distribution is *not* known a priori, and must be found by solving the master equation. Thus, it is impossible to discern if two different sets of rates will lead to the *same* NESS without solving both master equations. By contrast, for equilibrium systems with microscopically reversible dynamics, simply comparing the ratios of the rates will suffice!

In this letter, we address these fundamental issues. Starting from a *general* master equation which admits a unique stationary distribution, $P^*(\mathcal{C})$, we first review a graphical construction for P^* in terms of (directed) labelled trees. Established some time ago [3, 4, 5], this method seems not to be widely known. In this approach, the non-trivial *probability currents* associated with NESS are very naturally associated with the violation of detailed balance. Second, we propose a general classification of NESS in terms of *both* $P^*(\mathcal{C})$ and the (stationary) *probability currents*, $K^*(\mathcal{C}, \mathcal{C}')$, between configurations \mathcal{C} and \mathcal{C}' . In other words, we postulate that a complete description for a NESS is $\{P^*, K^*\}$, being the appropriate generalization of the Boltzmann distribution for equilibrium systems: $\{P^{eq}, 0\}$. In addition to the usual macroscopic averages $\langle A \rangle$, $\{P^*, K^*\}$ allows us to compute *fluxes* of all physical quantities (e.g., mass and energy currents), within and through our system. In this framework, we can specify the class of transition rates which lead to the same NESS, leading to a generalization of the “detailed balance condition” routinely exploited in simulation studies of equilibrium systems. In other words, all transformations of the rates which leave $\{P^*, K^*\}$ invariant are known. Further, *if* we are provided two distinct set of rates corresponding to the same NESS, then $\{P^*, K^*\}$ can be found trivially. We conclude with a discussion of entropy production and some general comments.

The master equation and graphic representation of P^ .* We begin with a generic master equation for an interacting many-body system with a finite number (N) of configurations. Labelling the configurations in some arbitrary fashion as $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_N$, we write the transition rate, per unit time, from \mathcal{C}_j to \mathcal{C}_i as w_i^j . All w ’s are real, non-negative, and assumed to be time-independent. In general, w_i^j differs from its reverse, w_j^i . The master equation for $P_i(t) \equiv P(\mathcal{C}_i, t)$, the probability to find the system in configuration \mathcal{C}_i at time t reads:

$$\partial_t P_i(t) = \sum_{j \neq i} [w_i^j P_j(t) - w_j^i P_i(t)] \equiv \sum_j W_i^j P_j(t) \quad (1)$$

i.e., the off- and on-diagonal elements of W are just w_i^j and $(-\sum_{j \neq i} w_i^j)$. Note that $\sum_i W_i^j = 0$, for $\forall j$, ensuring that $\sum_i P_i(t) = 1$, for $\forall t$. Such a W is known as a

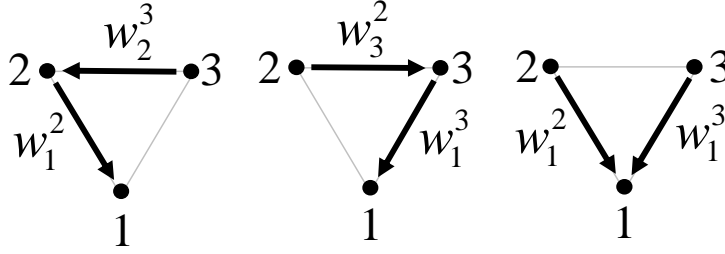


Figure 1. Representation of P_1^* , for a simple model with $N = 3$, in terms of directed labelled trees.

stochastic matrix. Since equation (1) is just a continuity equation for probability, we can write $K_i^j(t) \equiv w_i^j P_j(t) - w_j^i P_i(t)$ as the net *probability* current from \mathcal{C}_j into \mathcal{C}_i .

In the following, we assume that every configuration can be reached from every other configuration. Under these conditions, equation (1) is ergodic and has a unique stationary solution, $P_i^* \equiv \lim_{t \rightarrow \infty} P_i(t)$. The associated stationary currents are denoted by K_i^{*j} and satisfy $\sum_{j \neq i} K_i^{*j} = 0$, i.e., the total probability current into any given configuration vanishes. If the rates satisfy detailed balance, as for systems evolving towards thermal equilibrium where $w_i^j/w_j^i = P_i^{eq}/P_j^{eq}$, then all stationary currents *vanish*. An equivalent statement of detailed balance which does not reference P_i^{eq} explicitly [2] involves closed loops in configuration space, i.e., $\mathcal{C}_i \rightarrow \mathcal{C}_j \rightarrow \mathcal{C}_k \rightarrow \dots \rightarrow \mathcal{C}_n \rightarrow \mathcal{C}_i$. For each loop, we define the product of the associated rates in the “forward” and the “reverse” directions: $\Pi_f \equiv w_j^i w_k^j \dots w_n^i$ and $\Pi_r \equiv w_i^j w_j^k \dots w_n^i$. The detailed balance condition corresponds to $\Pi_f = \Pi_r$ for all loops. Related to integrability, this property allows P_i^* to be computed from the w ’s easily.

In the absence of detailed balance, P_i^* has to be found, in principle, from $\sum_j W_i^j P_j^* = 0$ laboriously. Fortunately, there is a systematic way to construct P^* [3, 4], using graph theoretical methods similar to those originally designed for electric networks [6]. First, associate each \mathcal{C}_i with a vertex, labelled i . Next, consider all distinct labelled trees (i.e., a graph consisting of all vertices with a single undirected edge between each pair, forming no loops). Denoting these trees as t_α , $\alpha = 1, 2, \dots, M$, Cayley’s theorem [7] states that $M = N^{N-2}$. To compute P_i^* , *direct every edge* towards the vertex i and denote this subset of (directed) trees by $t_{\alpha(i)}$. In other words, $\alpha(i)$ runs over the set of directed trees with i as the “root”. Next, a factor of w_k^n is associated with an edge directed from n to k . Finally, to each tree $t_{\alpha(i)}$, we assign a numerical value, $U(t_{\alpha(i)})$, which is the product of the $(N - 1)$ factors of w ’s in the tree. Clearly, $U(t_{\alpha(i)}) = 0$ if one of the associated rates vanishes. Then, the stationary distribution is given by

$$P_i^* = \mathcal{Z}^{-1} \sum_{\alpha(i)} U(t_{\alpha(i)}) \quad (2)$$

where \mathcal{Z} is just the normalization factor and may play the role of a (super-) partition function. We illustrate the procedure in figure 1 for $N = 3$. A more detailed discussion, including further examples, can be found in [8].

Probability currents and loops in configuration space. From the defining equation

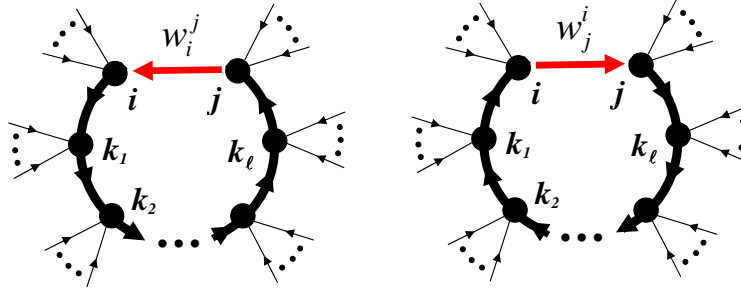


Figure 2. One term in [...] of equation (3), illustrating the emergence of nontrivial loops. Straight arrows (red online) represent the w factors; the rest (black online), the U 's.

for K_i^j above, we arrive at the *net* (stationary) probability current, from \mathcal{C}_j into \mathcal{C}_i :

$$K_i^{*j} \equiv w_i^j P_j^* - w_j^i P_i^* = \mathcal{Z}^{-1} \sum_{\alpha} \left[w_i^j U(t_{\alpha(j)}) - w_j^i U(t_{\alpha(i)}) \right] \quad (3)$$

Focusing on the expression within [...], we note that, for a specific α , the trees $t_{\alpha(i)}$ and $t_{\alpha(j)}$ differ only in the directed edges that connect vertices i and j (figure 2). Now, multiplication of $U(t_{\alpha(j)})$ by w_i^j can be regarded as adding a directed edge from j to i , converting $t_{\alpha(j)}$ into a graph with a *single* loop. Associated with this loop is the product $\Pi_i^j(t_{\alpha(j)}) \equiv w_i^j (w_{k_1}^i w_{k_2}^{k_1} \dots w_j^{k_\ell})$, where k_1, \dots, k_ℓ label the vertices between i to j . Similar considerations for $w_j^i U(t_{\alpha(i)})$ lead to a graph with the *same* loop, but traversed in the *opposite* sense and so, associated with $\Pi_j^i(t_{\alpha(i)}) = w_j^i (w_{k_\ell}^j \dots w_i^{k_1})$. Meanwhile, the rest of both trees (the side branches of the loops) are identical, so that $R(t_{\alpha(i)}) = R(t_{\alpha(j)})$, where R denotes the products of the w 's in the side branches. Summarizing, we write $w_j^i U(t_{\alpha(i)}) = \Pi_j^i(t_{\alpha(i)}) R(t_{\alpha(i)})$, etc., so that

$$K_i^{*j} = \mathcal{Z}^{-1} \sum_{\alpha} \left[\Pi_i^j(t_{\alpha(j)}) - \Pi_j^i(t_{\alpha(i)}) \right] R(t_{\alpha(i)}) \quad (4)$$

This expression explicitly demonstrates the emergence of nonzero steady state probability currents from rates which violate detailed balance, as manifested in irreversible loops.

A postulate. It is easy to construct cases where two sets of rates, one obeying detailed balance, and the other violating it, lead to the same stationary distribution. For example, for particles hopping on a ring with symmetric or biased rates, $P^{eq} = P^* \propto 1$. Hence, one might be tempted to consider any P^* as an “effective” equilibrium system. Indeed, nothing prevents us from labeling $-\ln P^*(\mathcal{C})$ as an “effective Hamiltonian”. Yet, in the simple example above, the system with biased rates carries a physical current while the system with symmetric rates does not. Thus, it is essential to identify a key signature that distinguishes a NESS from an equilibrium state. We believe that the nonvanishing stationary currents, K^* , fill this role. Therefore, we propose that $\{P^*, K^*\}$ form a *complete and unique* description for *any* stationary state. The major difference between a NESS and the standard equilibrium case is K^* being nonzero. In this sense, the class of NESS is significantly broader than equilibrium states, and their

analogues in electrodynamics would be, respectively, magnetostatics and electrostatics. Given $\{P^*, K^*\}$, all macroscopic stationary properties of the system can be computed. Beyond the usual averages with P^* as the weights, the K^* generate average *fluxes* (or currents) associated with physical observables, such as energy or particle number density. More details and examples will be discussed elsewhere [8].

Dynamic equivalence classes. A common notion of NESS is that, unlike their counterparts in thermal equilibrium, seemingly slight modifications of the rates lead to very different steady states. Armed with our characterization of a NESS in terms of $\{P^*, K^*\}$, we can pose a natural question: What determines the class of dynamics that leads to the same NESS? In other words, given a set of w 's and its associated NESS, what are the transformations (on the rates) which leave $\{P^*, K^*\}$ invariant? For the equilibrium case, $\{P^{eq}, 0\}$, the answer is well known: Any set of w 's will lead to a desired P^{eq} , provided they satisfy the “detailed balance condition”: $w_i^j/w_j^i = P_i^{eq}/P_j^{eq}$. This can be regarded as a constraint on the w 's, *given* a certain P^{eq} is to be achieved. In our framework, this constraint can now be easily generalized: To arrive at a *given* $\{P^*, K^*\}$ final state, the w 's must satisfy

$$w_i^j P_j^* - w_j^i P_i^* = K_i^{*j} . \quad (5)$$

for all pairs $i \neq j$. In the remainder of this letter, we will explore other representations of this constraint and some of its implications.

Let us decompose WP^* into its symmetric and antisymmetric parts:

$$W_i^j P_j^* = S_i^j + A_i^j \quad (6)$$

where $S \equiv (WP^* + WP^*)/2$, and $A \equiv (WP^* - WP^*)/2$. Then equation (5) is just the constraint that A is fixed to be $K^*/2$. In contrast, there is no such constraint on S , except for two restrictions. The physical rates must be non-negative ($w \geq 0$), leading to $S_i^j \geq |A_i^j|$, $\forall i \neq j$. Next, probability conservation imposes $\sum_i S_i^j = 0$, $\forall j$. Thus, we arrive at the conditions

$$S_i^j \geq \frac{1}{2} |K_i^{*j}| \quad \forall i \neq j, \quad S_j^j = - \sum_{i \neq j} S_i^j . \quad (7)$$

Within these constraints, we can choose *arbitrary* S 's and construct the associated transition rates via

$$W_i^j = \left[S_i^j + \frac{1}{2} K_i^{*j} \right] (P_j^*)^{-1} , \quad (8)$$

resting assured that the final NESS *will remain the same*. Thus, we may associate such S 's with an “equivalence class” of dynamical rates leading to the same given NESS.

It is very instructive to consider the difference, Δ_i^j , between two sets of rates that belong to the same class. Since the differences in the S 's are symmetric, Δ must satisfy

$$\Delta_i^j P_j^* = \Delta_j^i P_i^* . \quad (9)$$

Reminiscent of the ordinary detailed balance condition, it is possible to turn this into a mnemonic: “The *differences* (as opposed to the rates themselves) must satisfy detailed balance with respect to P^* .” Finally, we note an implication of this curious equation: If

two sets of rates are somehow known to generate the same NESS, their (non-vanishing) differences will provide a simple route to finding P^* . In contrast, for equilibrium states, the stationary distribution can be easily generated from just *one* set of rates, again due to the *known* absence of K^* 's in this case.

Entropy production. One of the key signatures of nonequilibrium steady states, recognized over three decades ago [9, 4, 10], is entropy production. For a general time-dependent solution of the master equation, two independent quantities were introduced: the entropy production of the “system” and of the “medium”,

$$\dot{\mathbf{S}}_{sys} \equiv \sum_{i,j} W_i^j P_j(t) \ln \frac{P_j(t)}{P_i(t)}, \quad \dot{\mathbf{S}}_{med} \equiv \sum_{i,j} W_i^j P_j(t) \ln \frac{W_i^j}{W_j^i}. \quad (10)$$

The former is readily recognized as the time derivative of $\mathbf{S}_{sys} \equiv -\sum_i P_i(t) \ln P_i(t)$, which motivates the term “entropy production of the system”. The latter is attributed to the coupling of the system to the *external environment* in a manner that prevents it from reaching equilibrium [4]. Neither $\dot{\mathbf{S}}_{sys}$ nor $\dot{\mathbf{S}}_{med}$ is necessarily positive. However, their sum, naturally termed the “total entropy production”

$$\dot{\mathbf{S}}_{tot} \equiv \sum_{i,j} W_i^j P_j(t) \ln \frac{W_i^j P_j(t)}{W_j^i P_i(t)}, \quad (11)$$

is indeed *non-negative* [4].

Recasting these expressions in terms of the probability currents [4], and taking $t \rightarrow \infty$ to focus on stationary states, we arrive at $\dot{\mathbf{S}}_{sys}^* = \frac{1}{2} \sum_{i,j} K^{*j}_i \ln (P_j^*/P_i^*)$ and $\dot{\mathbf{S}}_{med}^* = \frac{1}{2} \sum_{i,j} K^{*j}_i \ln (W_i^j/W_j^i)$. For equilibrium states, both trivially vanish since all currents are zero. By contrast, $K^* \neq 0$ for a NESS, though $\dot{\mathbf{S}}_{sys}^*$ remains zero (so that $\dot{\mathbf{S}}_{med}^* = \dot{\mathbf{S}}_{tot}^*$). The interpretation of these results is clear: In the steady state, the entropy “associated with our system” no longer changes. However, it is reasonable to expect that, being coupled in an irreversible way to the environment, such a NESS continues to “induce” the entropy of its surrounding medium to increase (indeed, $\dot{\mathbf{S}}_{med}^* = \dot{\mathbf{S}}_{tot}^* > 0$). In this sense, $\dot{\mathbf{S}}_{med}^*$ carries detailed information of transition rates and so, the precise nature of the coupling between our system and its environment. As a result, even if we insist on having the same NESS (i.e., a given $\{P^*, K^*\}$), $\dot{\mathbf{S}}_{med}^*$ will not be unique. Let us explore the implications of these “degrees of freedom”.

Since $\dot{\mathbf{S}}_{med}^* = \dot{\mathbf{S}}_{tot}^*$, we focus on the latter and exploit equation (6):

$$\dot{\mathbf{S}}_{tot}^* = \frac{1}{2} \sum_{i,j} K^{*j}_i \ln \frac{S_i^j + A_i^j}{S_i^j - A_i^j}. \quad (12)$$

Since $A = K^*/2$, the freedom we have is *any* S satisfying equation (7). An immediate consequence is that rates can be chosen to *minimize* the entropy production (associated with a given NESS), by having $S \gg A$. To lowest order in K^{*j}_i/S_i^j , we have $\dot{\mathbf{S}}_{tot}^* = \sum (K^*)^2 / (2S)$. Since $\dot{\mathbf{S}}_{med}^* \equiv 0$ for equilibrium cases, we can choose rates which are arbitrarily “equilibrium-like” in this respect. We should emphasize that, though $\dot{\mathbf{S}}_{tot}^*$ can be made arbitrarily small, it remains strictly positive and retains the NESS signature. At the opposite extreme, we can consider rates with “infinite

\dot{S}_{med}^* ” by lowering some S_i^j to $|A_i^j|$. Whether such a concept is useful deserves further exploration. It is natural to label such rates as “maximally asymmetric”, since one of the two directed edges between some pairs of configurations is *missing*. Such models abound in the literature, e.g., totally asymmetric exclusion processes (TASEP) [11]. One clear advantage of having maximally asymmetric rates for all edges is that the number of trees used for constructing P^* is kept at the absolute minimum. Of course, the expression for K^* also simplifies. Finally, the implications for non-equilibrium work theorems [12, 10] are not trivial and will be published elsewhere [8].

Conclusions. To summarize, we have addressed a fundamental question associated with non-equilibrium steady states: Within the framework of the master equation, what class, if any, of transition rates W lead to the same stationary state? For equilibrium systems, the answer is provided by the detailed balance condition. To generalize this answer to NESS, we first postulate that a NESS is completely and uniquely specified by its stationary distribution P^* *in conjunction with* the steady currents K^* . Then the *generalized* detailed balance condition is simply equation (5). Exploiting a graphic method to compute P^* in terms of directed labelled trees, we display the connection between K^* and “irreversible” loops - key characteristics of rates that violate detailed balance. Extensions, examples, and further implications of these explorations may be found in [8].

Acknowledgements. We thank U. Seifert for fruitful discussions. BS acknowledges the hospitality of the Isaac Newton Institute in Cambridge, UK, where some of this work was performed. Financial support from the NSF through DMR-0414122 is gratefully acknowledged.

- [1] Schmittmann B and Zia R K P 1995 Statistical mechanics of driven diffusive systems *Phase Transitions and Critical Phenomena* vol 17 ed C Domb and J L Lebowitz (New York: Academic)
- [2] Mukamel D 2000 Phase transitions in non-equilibrium systems *Soft and Fragile Matter: Nonequilibrium Dynamics, Metastability and Flow*, ed M E Cates and M R Evans (Bristol: Institute of Physics Publishing)
- [3] Hill T L 1966 *J. Theor. Biol.* **10** 399
- [4] Schnakenberg J 1976 *Rev. Mod. Phys.* **48** 571
- [5] Haken H 1983 *Synergetics, An Introduction* (Berlin: Springer, 3rd edition). See also Blythe R A 2001 (University of Edinburgh thesis: <http://www.ph.ed.ac.uk/cmatter/links/rab-thesis>).
- [6] Kirchhoff G 1847 *Poggendorff's Ann. Phys. Chem.* **72** 497
- [7] See, e.g., Roberts F S 1984 *Applied Combinatorics* (Englewood Cliffs: Prentice-Hall).
- [8] Schmittmann B and Zia R K P 2006 *to be published*
- [9] Glansdorff P and Prigogine I 1971 *Thermodynamic Theory of Structure, Stability, and Fluctuations* (London: Wiley)
- [10] Seifert U 2005 *Phys. Rev. Lett.* **95** 040602
- [11] For a review, see: Schütz, G M 2000 Exactly solvable models for many-body systems far from equilibrium *Phase Transitions and Critical Phenomena* vol 19 ed C Domb and J L Lebowitz (New York: Academic)
- [12] Jarzinsky C 1997 *Phys. Rev. Lett.* **78** 2690